

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



DEX-0279

BD

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
19 April 2001 (19.04.2001)

PCT

(10) International Publication Number  
WO 01/27276 A3

(51) International Patent Classification?: C12N 15/12. C07K 14/47, 16/30. A61K 38/17, 39/395. 48/00. G01N 33/53

(21) International Application Number: PCT/US00/28255

(22) International Filing Date: 11 October 2000 (11.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/417,031 12 October 1999 (12.10.1999) US

(71) Applicant (for all designated States except US): CORIXA CORPORATION [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): HARLOCKER, Susan, L. [US/US]; 7522 - 13th Avenue W., Seattle, WA 98117 (US). DILLON, Davin, C. [US/US]; 18112 NW Montreux Drive, Issaquah, WA 98027 (US). XU, Jiangchun [US/US]; 15805 SE 43rd Place, Bellevue, WA 98006 (US).

(74) Agents: POTTER, Jane, E., R.: Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
24 January 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/27276 A3

(54) Title: DNA SEQUENCES FROM BREAST TUMOR AND USES THEREOF

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Compositions may comprise one or more breast tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor protein, or mRNA encoding such a protein, in a sample are also provided.

## INTERNATIONAL SEARCH REPORT

Int'l. Appl. No.  
PCT/US 00/28255

A. CLASSIFICATION OF SUBJECT MATTER  
 IPC 7 C12N15/12 C07K14/47 C07K16/30 A61K38/17 A61K39/395  
 A61K48/00 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
 IPC 7 C12N C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	XU, J. ET AL.: "Identification of differentially expressed genes in human breast tumor using subtraction and microarray" PROC. AM. ASSOC. CANCER RES., vol. 40, March 1999 (1999-03), page 319 XP002160220 the whole document ---	1-49
A	JIANG, J. ET AL.: "Discovery of breast cancer antigens using subtracted cDNA libraries and cDNA microarrays" PROC. AM. ASSOC. CANCER RES., vol. 40, March 1999 (1999-03), page 32 XP002160221 the whole document --- -/-	1-49

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*&\* document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

23.05.2001

13 February 2001

Name and mailing address of the ISA  
 European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel: (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Mata-Vicente, M

# INTERNATIONAL SEARCH REPORT

Inte onal Application No  
PCT/US 00/28255

**C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HOUGHTON, R. L. ET AL.: "Real-time PCR and microarray analysis of breast cancer genes identified using differential display techniques" PROC. AM. ASSOC. CANCER RES., vol. 40, March 1999 (1999-03), pages 32-33, XP002160222 the whole document ----	1-49
A	WO 99 33869 A (CORIXA CORP) 8 July 1999 (1999-07-08) page 3 -page 4 claims 49-52 ----	1-49
A	YANG, G. P. ET AL.: "Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes" NUCLEIC ACIDS RES., vol. 27, no. 6, 15 March 1999 (1999-03-15), pages 1517-1523, XP002160223 page 1518, column 2, paragraph 3 page 1519, column 1, paragraph 2 ----	1,11
A	KONONEN, J., ET AL.: "Tissue microarrays for high-throughput molecular profiling of tumour specimens" NATURE MEDICINE, vol. 4, no. 7, July 1998 (1998-07), pages 844-847, XP002160224 page 844, column 2, paragraph 2 ----	

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/28255

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  
Although claims 19-23, 26, 29-31 and, as far as an "in vivo" method is concerned, claims 27 and 28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2.  Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

(1-49) - partially

#### Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1-49) - partially

Breast tumour-specific polypeptide encoded by polynucleotide SEQ ID NO:4 and fragments or variants thereof; antibodies against the same; antigen presenting cells expressing said polypeptide; vaccines, pharmaceutical compositions and diagnostic kits comprising any of the former. Uses thereof in diagnosis, prevention, monitorization and/or treatment of breast cancer.

2. Claims: (1-49) - partially

Idem as in subject 1, but limited to SEQ ID NO:9.

3. Claims: (1-49) - partially

Idem as in subject 1, but limited to SEQ ID NO:11.

4. Claims: (13-49) - partially

Breast tumour-specific polypeptide encoded by polynucleotide SEQ ID NO:1 and fragments or variants thereof; antibodies against the same; antigen presenting cells expressing said polypeptide; vaccines, pharmaceutical compositions and diagnostic kits comprising any of the former. Uses thereof in diagnosis, prevention, monitorization and/or treatment of breast cancer.

5. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:2.

6. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:3.

7. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:5.

8. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:6.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

9. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:7.

10. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:8.

11. Claims: (13-49) - partially

Idem as in subject 4, but limited to SEQ ID NO:10.

**INTERNATIONAL SEARCH REPORT**

## Information on patent family members

International Application No

PCT/US 00/28255

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9933869	A 08-07-1999	AU 2010699 A		19-07-1999
		EP 1042360 A		11-10-2000
		ZA 9811800 A		23-06-1999

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
19 April 2001 (19.04.2001)

PCT

(10) International Publication Number  
WO 01/27276 A2

(51) International Patent Classification<sup>7</sup>: C12N 15/12, C07K 14/47, 16/30, A61K 38/17, 39/395, 48/00, G01N 33/53

(74) Agents: POTTER, Jane, E., R.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(21) International Application Number: PCT/US00/28255

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 11 October 2000 (11.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/417,031 12 October 1999 (12.10.1999) US

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): CORIXA CORPORATION [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/27276 A2

(54) Title: BREAST TUMOR SEQUENCES AND METHODS OF USE THEREFOR

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as breast cancer, are disclosed. Compositions may comprise one or more breast tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a breast tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as breast cancer. Diagnostic methods based on detecting a breast tumor protein, or mRNA encoding such a protein, in a sample are also provided.

## BREAST TUMOR SEQUENCES AND METHODS OF USE THEREFOR

## TECHNICAL FIELD

The present invention relates generally to therapy and diagnosis of cancer, such as breast cancer. The invention is more specifically related to polypeptides comprising at least a portion of a breast tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of breast cancer, and for the diagnosis and monitoring of such cancers.

## BACKGROUND OF THE INVENTION

Breast cancer is a significant health problem for women in the United States and throughout the world. Although advances have been made in detection and treatment of the disease, breast cancer remains the second leading cause of cancer-related deaths in women, affecting more than 180,000 women in the United States each year. For women in North America, the life-time odds of getting breast cancer are now one in eight.

No vaccine or other universally successful method for the prevention or treatment of breast cancer is currently available. Management of the disease currently relies on a combination of early diagnosis (through routine breast screening procedures) and aggressive treatment, which may include one or more of a variety of treatments such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of treatment for a particular breast cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. *See, e.g.,* Porter-Jordan and Lippman, *Breast Cancer* 8:73-100, 1994. However, the use of established markers often leads to a result that is difficult to interpret, and the high mortality observed in breast cancer patients indicates that improvements are needed in the treatment, diagnosis and prevention of the disease.

Immunotherapies have the potential to substantially improve breast cancer treatment and survival. Such therapies may involve the generation or enhancement of an immune response to a breast tumor antigen. However, to date, relatively few breast tumor antigens are known and the generation of an immune response against such antigens has not been shown to be therapeutically beneficial.

Accordingly, there is a need in the art for improved methods for identifying breast tumor antigens and for using such antigens in the diagnosis and therapy of breast cancer. The present invention fulfills these needs and further provides other related advantages.

## SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as breast cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a breast tumor protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of sequences recited in SEQ ID NOS:1-11, variants of such sequences and complements of such sequences.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 9, preferably at least 15, amino acid residues of a breast tumor protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, vaccines are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and a non-specific immune response enhancer.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a breast tumor protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a non-specific immune response enhancer.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein or a polynucleotide encoding a fusion protein in combination with a non-specific immune response enhancer.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a breast tumor protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a breast tumor protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a breast tumor protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expresses such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be breast cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point

in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

#### DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the therapy and diagnosis of cancer, such as breast cancer. The compositions described herein may include breast tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (e.g., T cells). Polypeptides of the present invention generally comprise at least a portion (such as an immunogenic portion) of a breast tumor protein or a variant thereof. A "breast tumor protein" is a protein that is expressed in breast tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in a normal tissue, as determined using a representative assay provided herein. Certain breast tumor proteins react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with breast cancer. Polynucleotides of the subject invention generally comprise a DNA or RNA sequence that encodes all or a portion of a breast tumor polypeptide, or that is complementary to such a sequence. In general, a breast tumor-specific polynucleotide is expressed in breast tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in a normal tissue. Antibodies are generally immune system proteins, or antigen-binding fragments thereof, that are capable of binding to a polypeptide as described above. Antigen presenting cells include dendritic cells and macrophages that express a polypeptide as described above. T cells that may be employed within such compositions are generally T cells that are specific for a polypeptide as described above.

The present invention is based on the identification of expressed nucleotide sequences that are associated with breast tumors. Partial sequences of breast tumor-specific polynucleotides are provided in SEQ ID NOs:1-11.

#### BREAST TUMOR-SPECIFIC POLYNUCLEOTIDES

Breast tumor-specific polynucleotides generally comprise at least 15 consecutive nucleotides, preferably at least 27 consecutive nucleotides and more preferably at least 45 consecutive nucleotides, of a polynucleotide that is specifically expressed in breast tumor tissue. Alternatively, or in addition, such a polynucleotide may encode a breast tumor protein or a portion or other variant thereof as described herein. For example, a polynucleotide may encode at least nine consecutive amino acid residues of a breast tumor protein, preferably at least 15 consecutive amino acid residues. More preferably, a polynucleotide encodes an immunogenic portion of a breast tumor protein. Polynucleotides complementary to any such sequences are also encompassed by the present invention. Polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous breast tumor-associated polynucleotide or a portion thereof) or may comprise a variant of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native breast tumor protein or a portion thereof.

The percent identity for two polynucleotide or polypeptide sequences may be readily determined by comparing sequences using computer algorithms well known to those of ordinary skill in the art, such as Megalign, using default parameters. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, or 40 to about 50, in

which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Optimal alignment of sequences for comparison may be conducted, for example, using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. Preferably, the percentage of sequence identity is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the window may comprise additions or deletions (*i.e.*, gaps) of 20 % or less, usually 5 to 15 %, or 10 to 12%, relative to the reference sequence (which does not contain additions or deletions). The percent identity may be calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Variants may also, or alternatively, be substantially homologous to a native gene, or a portion or complement thereof. Such polynucleotide variants are capable of hybridizing under moderately stringent conditions to a naturally occurring breast tumor-associated DNA sequence (or a complementary sequence). Suitable moderately stringent conditions include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Polynucleotides may be prepared using any of a variety of techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening GenBank EST clones with mRNA from breast tumors and with mRNA from normal tissues. Such screens may be

performed, for example, using a Synteni (Palo Alto, CA) UniGem microarray according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Certain nucleic acid sequences of breast tumor-specific cDNA molecules identified by this process are provided in SEQ ID NOS:1-11. Alternatively, polypeptides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as breast tumor cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion may be used to isolate a full length gene from a suitable library (e.g., a breast tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with  $^{32}\text{P}$ ) using well known techniques. A bacterial or bacteriophage library is then screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences are then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to

the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence.

Polynucleotide variants may generally be prepared by any method known in the art, including chemical synthesis by, for example, solid phase phosphoramidite chemical synthesis. Modifications in a polynucleotide sequence may also be introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis (*see* Adelman et al., *DNA* 2:183, 1983). Alternatively, RNA molecules may be generated by *in vitro* or *in vivo* transcription of breast tumor-specific DNA sequences, or portions thereof, provided that the DNA is incorporated into a vector with a suitable RNA polymerase promoter (such as T7 or SP6). Certain portions may be used to prepare an encoded polypeptide, as described herein. In addition, or alternatively, a portion may be administered to a patient such that the encoded polypeptide is generated *in vivo* (*e.g.*, by transfecting antigen-presenting cells, such as dendritic cells, with a cDNA construct encoding a breast tumor polypeptide, and administering the transfected cells to the patient).



A portion of a sequence complementary to a coding sequence (*i.e.*, an antisense polynucleotide) may also be used as a probe or to modulate gene expression. cDNA constructs that can be transcribed into antisense RNA may also be introduced into cells of tissues to facilitate the production of antisense RNA. An antisense polynucleotide may be used, as described herein, to inhibit expression of a tumor protein. Antisense technology can be used to control gene expression through triple-helix formation, which compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules (*see* Gee et al., *In* Huber and Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co. (Mt. Kisco, NY; 1994)). Alternatively, an antisense molecule may be designed to hybridize with a control region of a gene (*e.g.*, promoter, enhancer or transcription initiation site), and block transcription of the gene; or to block translation by inhibiting binding of a transcript to ribosomes.

A portion of a coding sequence or of a complementary sequence may also be designed as a probe or primer to detect gene expression. Probes may be labeled with a variety of reporter groups, such as radionuclides and enzymes, and are preferably at least 10 nucleotides in length, more preferably at least 20 nucleotides in length and still more preferably at least 30 nucleotides in length. Primers, as noted above, are preferably 22-30 nucleotides in length.

Any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-, methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Nucleotide sequences as described herein may be joined to a variety of other nucleotide sequences using established recombinant DNA techniques. For example, a polynucleotide may be cloned into any of a variety of cloning vectors, including plasmids, phagemids, lambda phage derivatives and cosmids. Vectors of particular interest include expression vectors, replication vectors, probe generation vectors and sequencing vectors. In general, a vector will contain an origin of replication functional in at least one organism, convenient restriction endonuclease sites and one or more selectable markers. Other elements will depend upon the desired use, and will be apparent to those of ordinary skill in the art.

Within certain embodiments, polynucleotides may be formulated so as to permit entry into a cell of a mammal, and expression therein. Such formulations are particularly useful for therapeutic purposes, as described below. Those of ordinary skill in the art will appreciate that there

are many ways to achieve expression of a polynucleotide in a target cell, and any suitable method may be employed. For example, a polynucleotide may be incorporated into a viral vector such as, but not limited to, adenovirus, adeno-associated virus, retrovirus, or vaccinia or other pox virus (e.g., avian pox virus). Techniques for incorporating DNA into such vectors are well known to those of ordinary skill in the art. A retroviral vector may additionally transfer or incorporate a gene for a selectable marker (to aid in the identification or selection of transduced cells) and/or a targeting moiety, such as a gene that encodes a ligand for a receptor on a specific target cell, to render the vector target specific. Targeting may also be accomplished using an antibody, by methods known to those of ordinary skill in the art.

Other formulations for therapeutic purposes include colloidal dispersion systems, such as macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. A preferred colloidal system for use as a delivery vehicle *in vitro* and *in vivo* is a liposome (*i.e.*, an artificial membrane vesicle). The preparation and use of such systems is well known in the art.

#### BREAST TUMOR-SPECIFIC POLYPEPTIDES

Within the context of the present invention, polypeptides may comprise at least an immunogenic portion of a protein encoded by a breast tumor-specific polynucleotide (a breast tumor protein), or a variant thereof, as described herein. As noted above, a "breast tumor protein" is a protein that is expressed by breast tumor cells at a level that is at least two fold higher than the level in normal breast cells. Proteins that are breast tumor proteins may also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with breast cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties. Preferred polypeptides comprise at least a portion of a breast tumor protein that comprises a sequence encoded by any one of SEQ ID NOS:1-11, or a complement thereof.

An "immunogenic portion," as used herein, is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a breast tumor protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions

may contain a small N- and/or C-terminal deletion (e.g., 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (i.e., they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native breast tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (e.g., in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, <sup>125</sup>I-labeled Protein A.

As noted above, a composition may comprise a variant of a native breast tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (e.g., 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants preferably exhibit at least about 70%, more preferably at least about 90% and most preferably at least about 95% identity to the native polypeptide. The percent

identity may be determined as described above. Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein that co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast and higher eukaryotic cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems that secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity

matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be

chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided that comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably, the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, e.g., Stoute et al., *New Engl. J. Med.* 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza* B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (e.g., the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen present cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus*

*pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

#### BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a breast tumor protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a breast tumor protein if it reacts at a detectable level (within, for example, an ELISA) with a breast tumor protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about  $10^3$  L/mol. The binding constant maybe determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as breast cancer, using the representative assays provided herein. In other

words, antibodies or other binding agents that bind to a breast tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (e.g., blood, sera, leukophoresis, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (e.g., mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (i.e., reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained

from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include <sup>90</sup>Y, <sup>123</sup>I, <sup>125</sup>I, <sup>131</sup>I, <sup>186</sup>Re, <sup>188</sup>Re, <sup>211</sup>At, and <sup>212</sup>Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, *Shigella* toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (e.g., covalently bonded) to a suitable monoclonal antibody either directly or indirectly (e.g., via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulphydryl group, on one

may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulphydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group that is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by

encapsulation, such as within a liposome vesicle (e.g., U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

#### T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a breast tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the CEPRATE™ system, available from CellPro Inc., Bothell WA (see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a breast tumor polypeptide, polynucleotide encoding a breast tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a breast tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a breast tumor polypeptide if the T cells kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays

may be performed, for example, as described in Chen et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (e.g., by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a breast tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (e.g., TNF or IFN- $\gamma$ ) is indicative of T cell activation (see Coligan et al., *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a breast tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4 $^{+}$  and/or CD8 $^{+}$ . Breast tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, or from a related or unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4 $^{+}$  or CD8 $^{+}$  T cells that proliferate in response to a breast tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a breast tumor polypeptide (e.g., a short peptide corresponding to an immunogenic portion of such a polypeptide) with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a breast tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of a breast tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution. Following expansion, the cells may be administered back to the patient as described, for example, by Chang et al., *Crit. Rev. Oncol. Hematol.* 22:213, 1996.

#### PHARMACEUTICAL COMPOSITIONS AND VACCINES

Within certain aspects, polypeptides, polynucleotides, T cells and/or binding agents described herein may be incorporated into pharmaceutical compositions or immunogenic compositions (*i.e.*, vaccines). Pharmaceutical compositions comprise one or more such compounds and a physiologically acceptable carrier. Vaccines may comprise one or more such compounds and a non-specific immune response enhancer. A non-specific immune response enhancer may be any substance that enhances an immune response to an exogenous antigen. Examples of non-specific

immune response enhancers include adjuvants, biodegradable microspheres (e.g., polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition or vaccine.

A pharmaceutical composition or vaccine may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, that are efficiently transported into the cells.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide) and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of non-specific immune response enhancers may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (e.g., IFN- $\gamma$ , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an

administered antigen. In contrast, high levels of Th2-type cytokines (e.g., IL-4, IL-5, IL-6, IL-10 and TNF- $\beta$ ) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Ribi ImmunoChem Research Inc. (Hamilton, MT; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555. Another preferred adjuvant is a saponin, preferably QS21, which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprises an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210. Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient.

The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule or sponge that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane. Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. The amount of active compound contained within a

sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*) and based on the lack of differentiation markers of B cells (CD19 and CD20), T cells (CD3), monocytes (CD14) and natural killer cells (CD56), as determined using standard assays. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF $\alpha$  to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF $\alpha$ , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc<sub>Y</sub> receptor, mannose receptor and DEC-205 marker. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules (e.g., CD40, CD80 and CD86).

APCs may generally be transfected with a polynucleotide encoding a breast tumor protein (or portion or other variant thereof) such that the breast tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the breast tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

#### CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as breast cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed

using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8<sup>+</sup> cytotoxic T lymphocytes and CD4<sup>+</sup> T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that

cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (see, for example, Cheever et al., *Immunological Reviews* 157:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitory, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (e.g., more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 100 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (e.g., more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a breast tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated

using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

#### METHODS FOR DETECTING CANCER

In general, a cancer may be detected in a patient based on the presence of one or more breast tumor proteins and/or polynucleotides encoding such proteins in a biological sample (such as blood, sera, urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as breast cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a breast tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length breast tumor proteins and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10  $\mu$ g, and preferably about 100 ng to about 1  $\mu$ g, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (see, e.g., Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20<sup>TM</sup> (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with breast cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20<sup>TM</sup>. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as breast cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample

generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 $\mu$ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use breast tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such breast tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a breast tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient is incubated with a breast tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with a breast tumor polypeptide (e.g., 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of breast tumor polypeptide to serve as a control. For CD4<sup>+</sup> T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8<sup>+</sup> T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a breast tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a breast tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the breast tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a breast tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a

breast tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes that may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in any one of SEQ ID NOS:1-11. Techniques for both PCR based assays and hybridization assays are well known in the art (see, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample such as a biopsy tissue and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, breast tumor proteins and polynucleotides encoding such proteins may be used as markers for monitoring the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide detected by the binding agent increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple breast tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that result in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

#### DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a breast tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a breast tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a breast tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a breast tumor protein.

The following Example is offered by way of illustration and not by way of limitation.

## EXAMPLE

Identification of Breast Tumor Protein cDNAs

This Example illustrates the identification of cDNA molecules encoding breast tumor proteins.

Three prefabricated UniGem (UniGem 2.0-4.0) microarray chips (Synteni; Palo Alto, CA), each displaying 10,000 randomly selected GenBank EST cDNA clones from various tissues, were probed with mRNA pooled from three breast tumors and with mRNA pooled, in equal parts, from six normal essential tissues (heart, lung, liver, kidney, brain, colon). Forty-six clones were found to be over-expressed in the breast tumors and were subsequently purchased (GenomeSystems, St. Louis, MO) and further characterized using custom microarray technology to determine mRNA expression levels in a variety of tumor and normal tissues. Such analyses were performed using Synteni (Palo Alto, CA) custom microarrays, according to the manufacturer's instructions. Briefly, clones were arrayed on slides, with each product occupying a unique location in the array. The microarrays were probed with 1 mRNA probe pair (consisting of a normal tissue mRNA and a breast tumor mRNA) and the slides were scanned to measure fluorescence intensity. Data was analyzed using Synteni's provided GEMtools software. Of the 46 clones, eleven were shown to have at least two-fold over-expression in breast tumor tissue. The sequences of these 11 clones are provided herein as SEQ ID NOs:1-11, and are described in Table I.

Table I  
Breast Tumor-Associated Clones

Sequence	UniGem Expression Ratio	Microarray Expression Ratio	Sequence Homologies
15003 (SEQ ID NO:1)	12.3	3.2	Procollagen type I alpha-1 chain
15005 (SEQ ID NO:2)	7.0	2.7	Thrombospondin-2
15006 (SEQ ID NO:3)	6.4	4.0	Procollagen alpha 2
15769 (SEQ ID NO:6)	3.3	4.1	Human TREB protein
15774 (SEQ ID NO:7)	3.1	3.2	Human secreted protein gene 17
15780 (SEQ ID NO:8)	5.3	3.6	Human keratin 18

15791 (SEQ ID NO:9)	2.8	2.3	None
15796 (SEQ ID NO:10)	13.6	6.1	Human pregnancy-specific $\beta$ -1-glycoprotein
15957 (SEQ ID NO:4)	3.5	2.2	None
15962 (SEQ ID NO:11)	2.8	3.9	None
15964 (SEQ ID NO:5)	5.2	2.7	Human secreted protein gene 48

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

## CLAIMS

1. An isolated polypeptide comprising at least an immunogenic portion of a breast tumor-specific protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the breast tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:
  - (a) polynucleotides recited in any one of SEQ ID NOs:4, 9 or 11; and
  - (b) complements of the foregoing polynucleotides.
2. A pharmaceutical composition comprising:
  - (a) a breast tumor-specific polypeptide comprising at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:4, 9 or 11, or a complement of any of the foregoing sequences; and
  - (b) a physiologically acceptable carrier.
3. A vaccine comprising:
  - (a) a breast tumor-specific polypeptide comprising at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:4, 9 or 11, or a complement of any of the foregoing sequences; and
  - (b) a non-specific immune response enhancer.
4. A vaccine according to claim 3, wherein the non-specific immune response enhancer is an adjuvant.
5. A vaccine according to claim 3, wherein the non-specific immune response enhancer induces a predominantly Type I response.
6. A pharmaceutical composition comprising:
  - (a) a polynucleotide that encodes at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:4, 9 or 11, or a complement of any of the foregoing sequences; and
  - (b) a physiologically acceptable carrier.

7. A vaccine comprising:

- a) a polynucleotide that encodes at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOS:4, 9 or 11, or a complement of any of the foregoing sequences; and
- (b) a non-specific immune response enhancer.

8. A vaccine according to claim 7, wherein the non-specific immune response enhancer is an adjuvant.

9. A vaccine according to claim 7, wherein the non-specific immune response enhancer induces a predominantly Type I response.

10. A pharmaceutical composition comprising:

- (a) an antisense polynucleotide that is complementary to a polynucleotide that encodes at least 9 consecutive residues of a breast tumor-specific protein, wherein the breast tumor-specific protein is encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOS:4, 9 or 11, or a complement of any of the foregoing sequences; and
- (b) a physiologically acceptable carrier.

11. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a breast tumor protein that comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOS:4, 9 or 11 or a complement of any of the foregoing polynucleotide sequences.

12. A pharmaceutical composition comprising an antibody or fragment thereof according to claim 11, in combination with a physiologically acceptable carrier.

13. A pharmaceutical composition comprising:

- (a) an antigen-presenting cell that expresses a breast tumor-specific polypeptide comprising at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOS:1-11, or a complement of any of the foregoing sequences; and

(b) a pharmaceutically acceptable carrier.

14. A pharmaceutical composition according to claim 13, wherein the antigen presenting cell is a dendritic cell or a macrophage.

15. A vaccine comprising:

(a) an antigen-presenting cell that expresses a breast tumor-specific polypeptide comprising at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:1-11, or a complement of any of the foregoing sequences; and

(b) a non-specific immune response enhancer.

16. A vaccine according to claim 15, wherein the non-specific immune response enhancer is an adjuvant.

17. A vaccine according to claim 15, wherein the non-specific immune response enhancer induces a predominantly Type I response.

18. A vaccine according to claim 15, wherein the antigen-presenting cell is a dendritic cell.

19. A method for inhibiting the development of breast cancer in a patient, comprising administering to a patient an effective amount of a breast tumor-specific polypeptide comprising at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:1-11, or a complement of any of the foregoing sequences; and thereby inhibiting the development of breast cancer in the patient.

20. A method for inhibiting the development of breast cancer in a patient, comprising administering to a patient an effective amount of a breast tumor-specific polynucleotide that:

(a) encodes at least 9 consecutive residues of a breast tumor-specific protein encoded by a polynucleotide that comprises a sequence recited in any one of SEQ ID NOs:1-11, or a complement of any of the foregoing sequences; or

(b) is complementary to any sequence in (a);

and thereby inhibiting the development of breast cancer in the patient.

21. A method for inhibiting the development of breast cancer in a patient, comprising administering to a patient an effective amount of an antibody or antigen-binding fragment thereof that specifically binds to a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences, and thereby inhibiting the development of breast cancer in the patient.

22. A method for inhibiting the development of breast cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least 9 consecutive amino acids of a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences, and thereby inhibiting the development of breast cancer in the patient.

23. A method according to claim 22, wherein the antigen-presenting cell is a dendritic cell.

24. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (i) polynucleotides recited in any one of SEQ ID NOs:1-11; and
- (ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

25. A method according to claim 24, wherein the biological sample is blood or a fraction thereof.

26. A method for inhibiting the development of breast cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 24.

27. A method for stimulating and/or expanding T cells specific for a breast tumor protein, comprising contacting T cells with one or more of:

(i) a polypeptide comprising at least 9 consecutive amino acids of a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences;

(ii) a polynucleotide encoding such a polypeptide; and/or

(iii) an antigen presenting cell that expresses such a polypeptide;

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

28. An isolated T cell population, comprising T cells prepared according to the method of claim 27.

29. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 27.

30. A method for inhibiting the development of breast cancer in a patient, comprising the steps of:

(a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with at least one component selected from the group consisting of:

(i) a polypeptide comprising at least 9 consecutive amino acids of a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences;

(ii) a polynucleotide encoding such a polypeptide; and

- (iii) an antigen-presenting cell that expresses such a polypeptide; such that T cells proliferate; and
- (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of breast cancer in the patient.

31. A method for inhibiting the development of breast cancer in a patient, comprising the steps of:

- (a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with at least one component selected from the group consisting of:
  - (i) a polypeptide comprising at least 9 consecutive amino acids of a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences;
  - (ii) a polynucleotide encoding such a polypeptide; and
  - (iii) an antigen-presenting cell that expresses such a polypeptide; such that T cells proliferate;
- (b) cloning at least one proliferated cell; and
- (c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of breast cancer in the patient.

32. A method for determining the presence or absence of breast cancer in a patient, comprising the steps of:

- (a) contacting a biological sample obtained from a patient with a binding agent that binds to a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:
  - (i) polynucleotides recited in any one of SEQ ID NOs:1-11; and
  - (ii) complements of the foregoing polynucleotides;
- (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and
- (c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of breast cancer in the patient.

33. A method according to claim 32, wherein the binding agent is an antibody.

34. A method according to claim 33, wherein the antibody is a monoclonal antibody.

35. A method for monitoring the progression of breast cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of breast cancer in the patient.

36. A method according to claim 35, wherein the binding agent is an antibody.

37. A method according to claim 36, wherein the antibody is a monoclonal antibody.

38. A method for determining the presence or absence of breast cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of breast cancer in the patient.

39. A method according to claim 38, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

40. A method according to claim 38, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

41. A method for monitoring the progression of breast cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotides;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of breast cancer in the patient.

42. A method according to claim 41, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

43. A method according to claim 41, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

44. A diagnostic kit, comprising:

(a) one or more antibodies, or antigen-binding fragments thereof, that specifically bind to a breast tumor-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOs:1-11 or a complement of any of the foregoing polynucleotide sequences; and

(b) a detection reagent comprising a reporter group.

45. A kit according to claim 44, wherein the antibodies are immobilized on a solid support.

46. A kit according to claim 45, wherein the solid support comprises nitrocellulose, latex or a plastic material.

47. A kit according to claim 44, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

48. A kit according to claim 44, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

49. A diagnostic kit, comprising:

(a) an oligonucleotide that specifically hybridizes to a polynucleotide that encodes a breast tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NOS:1-11 or a complement of any of the foregoing polynucleotides; and

(b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

## SEQUENCE LISTING

<110> Corixa Corporation  
Harlocker, Susan L.  
Dillon, Davin C.  
Xu, Jiangchun

<120> BREAST TUMOR SEQUENCES AND METHODS OF  
USE THEREFOR

<130> 210121.486PC

<140> PCT  
<141> 2000-10-11

<160> 11

<170> FastSEQ for Windows Version 3.0

<210> 1  
<211> 720  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(720)  
<223> n = A,T,C or G

<400> 1  
ggcacgagat caatgctggc agccagtttg aatataatgt agaaggagtg acttccaagg 60  
aatggctac ccaacttgcc ttcatgcgcc tgctggccaa ctatgcctct cagaacatca 120  
cctaccactg caagaacagc attgcataca tggatgagga gactggcaac ctgaaaaagg 180  
ctgtcattct acagggctct aatgatgttgc aacttggttgc tgagggcaac agcaggttca 240  
cttacactgt tctttagat ggctgctcta aaaagacaaa tgaatggggg aaagacaatc 300  
attgaataca aaacaaataa gccatcacgg cctgccttc cttgatattt cacctttgga 360  
catcggtgt gctgaccagg aattctttgt ggacattggc ccagtctgtt tcaaataaat 420  
gaactcaatc taaattaaaa aagaaagaaa ttgaaaaaaa ctttctctt gccatttctt 480  
cttcttctt ttaactgaa agcntgaatc cttccatttt cttctggcnc atcttacctt 540  
ggctttaaa ntntgggnnn nagnaaaaaa naaggantga tcagagcatt gtgcaataca 600  
gttccattaa ctccctcccc cgctccccca aaaatttgc aaaaaatttc aacactctt 660  
cacctgttat gaaaaatgtc aacatttgta agaaaaccaa aataaaaaaa aaaaaaaaaa 720

<210> 2  
<211> 493  
<212> DNA  
<213> Homo sapien

<400> 2  
taccgattaa ctaatcgaa catgtaaaac atgttacaaa aataaacgaa ctctcctttt 60  
gtcctacaat gaaagccctc atgtgcagta gagatgcagt ttcatcaaag aacaaacatc 120  
cttgc当地 atggtgtgacgc ggttccagat gtggatttgg caaaacctca ttttaagtaaa 180  
aggttagcag agcaaagtgc ggtgc当地 tag ctgtgc当地 tgccgctgtg gcgtc当地 240  
ggctc当地 tgagcttct tccccagctt tgctgc当地 gaggaaccag agcagacgca 300

caggccggaa aaggcgcatc taacgcgtat ctaggcttg gtaactgcgg acaagttgct	360
tttacctgat ttgatgatac atttcattaa ggttccagtt ataaatattt tggtaatattt	420
tattaagtgg actatagaat gcaactccat ttaccagtaa cttatttaa atatgcctag	480
taaccacata tgg	493

<210> 3  
 <211> 736  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(736)  
 <223> n = A,T,C or G

<400> 3	
gatttgcctt cgaggccana attcggcacg aggcttaca aacatatgtc caaggactct	60
aaattgagac tcttccacat gtacaatctc atcatcctga antctataat gaagaaaaaag	120
atctagaac tgagttgngg agctgactct aatcaaatgt gatgattgga attagaccat	180
ttggcccttg aactttcata ggaaaaatga cccaacattt cttagcatga gctacccat	240
ctctagaagc tgggatggac ttactattct tggatataattt ttagatactg aaaggtgcta	300
tgcttcgtt attattccaa gactggagat aggccaggct aaaaaggtat tattatttt	360
ccttaatga tggtgctaaa attcttccta taaaattccct taaaataaaa gatggttaa	420
tcactaccat tggaaaaaca taactgttag acttccgtt tctgaaaagaa agagcatcgt	480
tccaatgtt gttcnctgtt cctctgtcat actgtatctg gaatgcttgc taataacttgc	540
atgcttccta gaccagaaca tggtaggtccc cttgtgtctc aagactttt ttttcttaat	600
tgcatttgcattt ggctctattt taattttttt cttttaaaaat aaacagctgg gaccatccca	660
aaagacaagc catgcataca actttggtca tggatctctg caaagcatca aattaaatgc	720
acgctttgtt catgtc	736

<210> 4  
 <211> 649  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(649)  
 <223> n = A,T,C or G

<400> 4	
gaaagactga gttgctcagg cctaggctta gaatttgctg cgtttggaa ataaaaagaac	60
aaaatgatac attagcctgc catatcaaaa acatataaaa gagaaattat ccctaagtca	120
aggccccca taagaataaa atttcttatt aaggcattt gatgtcatgt aatccctttc	180
aaagtgcagt atgaaaacaa agggaaaaac actgaagcac acgcaactct cacagcgaca	240
ttttctgacc cacgaatgtat gccttgggtg ggcaacacga ttgcattgtt tggagacact	300
tcggaagtaa atgtggatga gggaggagct gtccttgcattt tggatgtt gacattacag	360
atacctcctc ttgaagaagg aataataatgtt ttaatcaaaa aagaagacta aaaaatgtaa	420
aatttggaaag gaatccataa atgcgtgtt gtctaaataac aaattatcat gtgaagaaaa	480
ggcccaagtgc tccaataag cagaccttgc ttggatgtt ggctaaattat gaatggggaa	540
tactggccag ttaatttcca gttttaatgg aaaacngatc caagaagaaa ttatgttgaat	600
aggtaangg ctgggctttg aggcttataa cctanaaaagg atgggtggg	649

<210> 5  
 <211> 550

<212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(550)  
 <223> n = A,T,C or G

<400> 5

cgcgccctgca	gacacagcat	ctactcagcg	tgggtcacct	ctgtgaacat	cactgactgc	60
aaggcctccct	caatttctgg	tgcagcccat	cagggaccca	cagcgccctgg	gaggatggtg	120
cgatcttgg	ccaatgggga	aatcgtgcag	gacgacgacc	cccgagttag	gaccactacc	180
cagccaccaa	gaggtagcat	tcctcgacag	agcttcttca	ataggggcca	tggtgetccc	240
ccagggggtc	ctggcccccgg	ccagcagcag	gcaggtgcca	ggctgggtgc	tgctcaagtcc	300
cccttcaatg	acctcaaccg	gcagctggtg	aacatgggct	ttccgcagtg	gcacatctggc	360
aaccatgtt	gtggagccgg	tgacctccat	cctgctctct	tcctgctcat	gatgcttggt	420
ggtcgtggcc	tncttctgggt	tggccttgtc	tacctggtggt	cccacctgaa	tcaagcggtg	480
accttttggaa	gggctgatan	gggggtggggt	ttgttggaaa	ggaacttgc	gggcctnggt	540
ggggaaaaaca						550

<210> 6  
 <211> 679  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(679)  
 <223> n = A,T,C or G

<400> 6

cctcgtgccc	attcttggga	ggacactttt	gccaatgaac	tctttcccca	gctgatttagt	60
gtctaaggaa	tgttcaata	ctgttgcctt	tttccttgac	tattacactg	cctggaggat	120
agcagagaag	cctgtctgt	tttcattcaa	aaagccaaaa	tagagagtagt	acagtccctag	180
agaattcctc	tatttgtca	gatctcatag	atgaccccca	ggtattgtct	tttgacatcc	240
agcagtccaa	ggtattgaga	catattactg	gaagtaagaa	atattactat	aattgagaac	300
tacagcttt	aagattgtac	ttttatctta	aaagggtggt	agttttccct	aaaatactta	360
ttatgttaagg	gtcatttagac	aaatgtctt	aaatgtacat	ggaattttatg	aatggttctt	420
tatcatttct	cttccccctt	tttggcatcc	tggcttgcc	ccagttttag	gtccttttagt	480
ttgctctgt	aagcaacggg	aacacctgct	gagggggtct	tttcccttca	tggatacttc	540
aagtaagatc	aagaatcttt	tgtgaattta	tagaaattta	ctatgtaaat	gcttgcgttga	600
attttttctt	gctagtgnag	ttcngaaag	ggctttctcc	atttatttaa	aactacccat	660
gcatttaaaa	ggncatgc					679

<210> 7  
 <211> 1177  
 <212> DNA  
 <213> Homo sapien

<400> 7

gcatttgtctg	tgggtgtgacc	atagcagatt	atatttggtt	cctgaatgtt	tgtggtgctt	60
atttctgtgt	ttgttccaag	cggttcaagtc	atgcccattgc	ctgcctcggt	agatggagta	120
atgtacaatg	aactccatga	gtctctccag	ggctgcctgc	agcacgtttt	ttccaagtag	180
cctatttggaa	ttcccatctc	aatgtcttg	gatgcgagcg	tcagcggctc	cagagctcggt	240
ggcggttgag	gtcccccttgg	gggaaccctt	tcctggccat	cgaggtcggt	gggctgcccgt	300

ctgtgggcag	gaggacccga	ggggcagcca	ggaaaggcga	tctcttca	ctgtaaaagtt	360
gccccgggtgc	agccgccttt	ccttctacca	tggaaaatgc	aggctggggc	cttgggggtga	420
gcctgcgggg	ctctgggtct	gtccccgacc	cccaccacca	ccagaatgca	gttccagctt	480
aggaagccac	aaacaagcca	cccaggagga	acaaaacacc	gccagcgtgg	attttccaaa	540
tttcccttgg	aagtaagtct	cgctcttgcc	aaagaaaaagt	ctggcttgg	gagtctctgg	600
agcccaggat	gccagcatgt	gccaatgact	gtcaccttca	tctcttcaaa	agaaaaagcca	660
tagccgagga	ctgtcccgcg	accccccgtgg	actgcgtcta	ggtcatgtga	ttctgttttc	720
atttctcatc	ccatccaatt	tgtccctttc	tcctgtcatt	ttcttcctct	gtggttccctt	780
caaagtgtt	ataatttgt	ctgaacttca	aaatgtgtcc	cgttctcccc	agaccactct	840
agccacagta	tatttcaata	aaattacttc	ttatatttgc	agaaaattctt	ttgggtgtat	900
tttatttttt	cctctcaata	tatataatttgc	gacaaacgct	ggcaaaaaga	aaaaaatggt	960
aagcaaaaaa	cccaagataa	agtttcgagg	acatcaggcc	ttttgaaata	caatgtcaaa	1020
tgacacattt	tacgttttca	aaaaatccgc	tagacatgtc	ataagtttt	actgtaatgc	1080
ccagggaaagg	atatcttaaa	atattctaaa	cttgggtgtac	aaagggataaa	ttaactgtaa	1140
cagttttca	ataaaatcgag	ttgggggttt	ccaccgt			1177

<210> 8  
<211> 396  
<212> DNA  
<213> *Homo sapien*

```
<220>
<221> misc_feature
<222> (1)...(396)
<223> n = A,T,C or G
```

<400> 8

gggggcacca	cagtggcac	cacacagtc	gctgagg	tg	gagctgctg	gacgcac	gctc	60
acagagctga	gacgtacagt	ccagtcctg	gagatgc	acc	tggactccat	gaga	aaatctg	120
gtgagtgcct	tcacatcacc	tgcccag	ctc	cttca	tgcaga	ccca	accctg	180
tctcaaccca	aatccatacc	ctcatatcat	gagt	tcctt	agctc	agaaa	gagt	240
ccttttgc	tttccctcca	ctcctatccc	ttatccc	agt	acttggcaca	tagc	agggtgc	300
ccaaaaaaagt	ttccaaaaagt	gaagggatg	ta	gcagccctgg	gactctggc	tcacc	ctgccc	360
cctcctctct	gtccccctgc	angaaggcccc	agctt	g				396

<210> 9  
<211> 636  
<212> DNA  
<213> Homo sapien

```
<220>
<221> misc_feature
<222> (1)...(636)
<223> n = A.T.C or G
```

<400> 9

```

agcttgttgc cgcactttt tttttttttt ttttgaatta aaataacaata ttttattttt 60
gtcacctaaa aaacaagcaa cccagantta agatatacgt aagtttctca aatttatcct 120
aataaaattt tgaatattta tacaatttct acagaaaacat acagngnac 180
taaatcaaat ttataaaatg ttagaaaaatg catagtgata ttaccaactt acaaagcaag 240
gatatggaa ttatttccaa agcagccctac agtagaaaaat agtcattatg gcagcagctt 300
ctgatgttt tggggatggtag gttttctgat ttcaatatat agaatcatat tcataagata 360
tcttcttta acgaatttgc caaagtaccc attttaaaatt tacatgcaca gttcatttgc 420
acctttctta ggccttatgca tagtttataaa ggttataatc tactcaacat ggaaaatgga 480
gccttatttgc aaacacacaaa gtaattttaaag taccatttctt ctcttagttt ctttttttat 540

```

agttggttta ttttgcattt ataaatgtta aacatcccta gagatgaaag ttaaaaatggt	600
tgatcacaga tcagtagcaa aatacaaatt gacaat	636
<210> 10	
<211> 885	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(885)	
<223> n = A,T,C or G	
<400> 10	
ttttttagaa aacaactgtg tttaatgaca atcctcagtt taaaaaaaaa aaagaagaag	60
aagggttata cacactgtac acacattac aatggcttg gaggatagca gtgctgcgaa	120
aagggttca ggaggatccg gcctggaca ggattgaggt atgttcagc ctccaggccc	180
tgggtctcc tgcataaga atacccctcc ccatttgcact gtgaactttt tggcctggat	240
tctggagaac agatttccag gattgtcagc cagaaggcag acagatgcag gcacccatca	300
agacctgacc tcaggaagtg gccctgcct acagcccagt tgctcagcca gggctgaagg	360
ccatggggcc ccagcacccct tgcttcagtg ccagccctg gaaggaacct cacaacaggg	420
atacagcaag gacactccag tccccccagt cctgccccatgg tgctaccctg agggacagg	480
atggagacag ggcagccagg tttgccccatgg cctgcatacg gggcccaaga ctgccttcc	540
tcttaagtca tgccaaagcc tccctgcctca gtctgagaca gtgcgtggca ggtgaccacg	600
acctgcgtgg ccctcccccggc agttgtcatg gtgttgtac cccacccat cccctgagg	660
agacatgggc tcagtcacat gcctggccat gcctggcc cacagccaca aagatggca tgggtctcta	720
gcctgatatt cgtggccctgg caggggtcag caacccctga nggcatccaa nccatggtca	780
nangaaagtg tnggcaggct nggcacagcc aaagaagtca ggacccacga gacggggaa	840
gccttccaga gccttcaccc tcacagggtc aaacttccag tagag	885
<210> 11	
<211> 766	
<212> DNA	
<213> Homo sapien	
<220>	
<221> misc_feature	
<222> (1)...(766)	
<223> n = A,T,C or G	
<400> 11	
tttcccaaca tatgaactgc agcataatcac tttttttttt tgcctcag gttcctcagc	60
tgtaaaaattt aaaaatataat gtatataataa tattataat aataataat gtaatgttagt	120
acttgggggtt aaagcacttt gagatcctt gttgaaaggc accataggag tgccaaatgt	180
tattatgtgg ccaagggggg tatttaact gtcagttccc aaaggccagg aaaggttggg	240
gtcattttttn ttaaagacga gctgtaaaata tcaacttaggc agccaaatgt gttgactatg	300
aagatgcaaa actattacta ggctgataaa atcatagttt cttaatggct accaataagg	360
caaataatcnc aataataaaac gccaaattcc ttagggcgga ctatggaca accacatgg	420
aaactttggg ggaggcatga ggggggaaca tctcaaaatg ccaatgtaaa atttaactta	480
cagcaatatt caccagcaga aaatgtctt catatggaaat gatttcatgt tgctaaagaaa	540
aagaattcaa tttgtgtcc tgatgttcaat actagaatgt tggctataat agttctgttc	600
ttacaacaca tggaaatttt tggatgttcaat actagaatgt tggctataat agttctgttc	660
ctactcacaa acatgttctt ggtgtatttc ttatgcaac aatcttcagg cagcaaagat	720
gtctgttaca tctaaacttg aataataaaag ttttaccncc agttac	766